Hierarchical Content-Based Image Retrieval of Skin Lesions

Hierarchical Content-Based Image Retrieval of Skin Lesions

Lucia Ballerini[1], Robert B. Fisher[1], Ben Aldridge[2], and Jonathan Rees[2]

[1] School of Informatics, University of Edinburgh, UK
lucia.ballerini@ed.ac.uk, rbf@inf.ed.ac.uk
[2] Dermatology, University of Edinburgh, UK
ben.aldridge@ed.ac.uk, jonathan.rees@ed.ac.uk

Abstract. This paper proposes a novel hierarchical content-based image retrieval system and its application to skin lesion images. Five common classes of skin lesions, including two non-melanoma cancer types, are used. Colour and texture features are extracted from lesions. Feature selection is embedded in a hierarchical framework that chooses the most relevant feature subsets by comparing different similarity for each level of the hierarchy. Experiments on our database of 533 images show that the proposed hierarchical scheme improves retrieval precision by about 4%, reaching a maximum average precision of 78%.

1 Introduction

Research in content-based image retrieval (CBIR) today is an extremely active discipline. There are already review articles containing references to a large number of systems and description of the technology implemented [1, 2]. A more recent review [3] reports a tremendous growth in publications on this topic. Applications of CBIR systems to medical domains already exist [4], although most of the systems currently available are based on radiological images. A query-by-example CBIR involves providing the CBIR system with an example image and retrieves the most visually similar images. This is our goal as described later.

Feature selection is one of the key challenges for optimisation of CBIR system. However, assessment of feature performance and feature selection methods in CBIR have to be carried out in a slightly different ways from classification and categorisation [5]. From the image processing point of view, it is important to gather as much features as possible to represent the images, yielding vectors with hundreds or even thousands of features to represent the images. However, the large number of features actually represents a problem. It leads to the "dimensionality curse" problem [6], where the indexing structures degrade and the significance of each feature decreases, making the process of storing, indexing and retrieving extremely time consuming. Moreover, in several situations, many features are correlated, meaning that they bring redundant information about the images that can deteriorate the ability of the system to correctly distinguish them. To avoid this problem, feature selection techniques can be employed to

reduce the feature vector size. Another problem is the "semantic gap", where the low-level features automatically extracted from images do not satisfactorily represent the semantic interpretation of the images. In fact, several challenges in CBIR systems are still opened and researchers are endeavouring to solve them, e.g. "what features best represent a given set of images?", and "what distance function most approximates the human perception of similarity among the images of a given dataset?" [7]. We try to address these challenges with the proposed hierarchical system.

Hierarchical retrieval schemes have been proposed since very long time [8, 9]. However, there are few approaches that incorporate a hierarchical design in CBIR system. Chow et al. [10] proposed a tree-structured image representation, where a root node contains the global features, while child nodes contain the local region-based features. A multilayer self-organising map is used to process the tree structured image data. Global and local features are also used in the Image Retrieval in Medical Applications (IRMA) approach described by Lehmann et al. [11]. The partitioning in IRMA is computed in the image domain, using a blob representation, where image regions are approximated by their best fitting ellipses. This partitioning permit a hierarchical image decomposition to model a multiscale approach. A hierarchical subspace method is presented by Wichert [12]. In their method, the subspaces correspond to different resolution of the images. During image retrieval a coarse to fine search at different resolutions is carried out.

Forming a hierarchy of features for retrieval has been explored by other researchers, but their goals for doing so differ from ours [13]. For example, Dy et al. [13] describe a "customized-queries" approach. Their approachfirst classifies the query to one labelled class and then customises the query to that classes by using the features that best distinguish the subclasses within the chosen class. Swets et al. [14] proposed a Self-Organizing Hierarchical Optimal Subspace Learning and Inference Framework. It uses the theories of linear discriminant projection for automatic optimal feature selection in each of the internal nodes of a Space-Tessellation Tree. In the framework, principal component analysis is to produce a set of most expressive features, and subsequently linear discriminant analysis is to produce a set of most discriminant features.

On the other hand, a large number of combination of classifiers in a hierarchical structure have been proposed in the literature and their improvement in the recognition accuracy have been shown in a large number of experimental studies [6]. For these reasons we believe that CBIR systems can also benefit from a hierarchical approach that combines retrieval sub-systems.

Our approach substantially differs from the above hierarchical CBIR proposals. Our system uses different feature sets for comparing similarity at each level of the hierarchy, and for each group of classes. The hierarchy is fixed a priori during the learning phase. In the course of the operative phase the system automatically discovers the best feature subset for each query image without the need of classifying it in a given class. Our approach is not limited in applicability to medical domains, but can be applied to any domain where the features that

best discriminate some given classes are different from those that characterise other classes.

In this paper we describe and evaluate an application of the hierarchical CBIR system to skin lesion images.

Dermatology atlases containing a large number of images are available online [15, 16]. However, their searching tool only allows query by the name of the lesion. On the other hand, the possibility of retrieving images based on visual similarity would greatly benefit both the non-expert users and the dermatologists. As already pointed out [4, 17], there is a need for CBIR as a decision support tool for dermatologists in the form of a display of relevant past cases, along with proven pathology and other suitable information. CBIR could be used to present cases that are not only similar in diagnosis, but also similar in appearance within the same class, and cases with visual similarity but different diagnoses. Hence, it would be useful as a training tool for medical students and researchers to browse and search large collection of disease related illustrations using their visual attributes.

To our knowledge, there are few CBIR systems in dermatology. Chung et al. [18] created a skin cancer database. Users can query the database by feature attribute values (shape and texture), or by synthesised image colours. It does not include a query-by-example method, as do most common CBIR systems. Their report concentrates on the description of the web-based browsing and data mining. However, nothing is said about database details (number, lesion types, acquisition technique), nor about the performance of the retrieval system. Celebi et al. [19] developed a system for retrieving skin lesion images based on shape similarity. The novelty of that system is the incorporation of human perception models in the similarity function. Results on 184 skin lesion images show significant agreement between computer assessment and human perception. However, they only focus on silhouette shape similarity and do not include many features (colour and texture) described in other papers by the same authors [20]. Rahman et al. [17] presented a CBIR system for dermatoscopic images. Their approach include image pre-processing, segmentation, feature extraction (colour and textures) and similarity matching. Experiments on 358 images of pigmented skin lesions from three categories (benign, dysplastic nevi and melanoma) are performed. A quantitative evaluation based on the precision curve shows the effectiveness of their system to retrieve visually similar lesions (average precision $\simeq 60\%$). Dorileo et al. [21] presented a CBIR system for wound images (necrotic tissue, fibrin, granulation and mixed tissue). Features based on histogram and multispectral co-occurrence matrices are used to retrieve similar images. The performance is evaluated based on measurements of precision ($\simeq 50\%$) on a database of 215 images. All these approaches only consider a few classes of lesions and/or do not exploit many useful features in this context.

Most of the work in dermatology has focused on skin cancer detection. Different techniques for segmentation, feature extraction and classification have been reported by several authors. Concerning segmentation, Celebi et al. [22] presented a systematic overview of recent border detection methods: clustering

followed by active contours are the most popular. Numerous features have been extracted from skin images, including shape, colour, texture and border properties [23–25]. Classification methods range from discriminant analysis to neural networks and support vector machines [26, 27, 20]. See Maglogiannis et al [28] for a review of the state of the art of computer vision system for skin lesion characterisation. These methods are mainly developed for images acquired by epiluminescence microscopy (ELM or dermoscopy) and they focus on melanoma, which is actually a rather rare, but quite dangerous, condition whereas other skin cancers are much more common.

Motivated by this, we propose a hierarchical retrieval system that focuses on 5 common classes of skin lesions: Actinic Keratosis (AK), Basal Cell Carcinoma (BCC), Melanocytic Nevus / Mole (ML), Squamous Cell Carcinoma (SCC), Seborrhoeic Keratosis (SK). The present work extends our previous work on CBIR [29, 30]. The main difference of this work to respect to our previous published works is the proposal of a novel retrieval method based on a hierarchical structure and a feature selection merging scheme.

The paper is organised as follow: Section 2 describes the colour and texture features. Section 3 defines the similarity criteria. Section 4 is devoted our new proposal. Results are presented in Section 5. Conclusions follow.

2 Feature extraction

CBIR requires the extraction of several features from each image, which, consequently, are used for computing similarity between images during the retrieval procedure. These features describe the content of the image and that is why they must be appropriately selected according to the context. The features have to be discriminative and sufficient for the description of different pathologies. Basically, the key to attaining a successful retrieval system is to choose the right features that represent each class of images as uniquely as possible.

Many feature extraction strategies have been proposed [23, 24] from the perspective of classification of images as malignant or benign. Different features attempt to reflect the parameters used in medical diagnosis, such as the *ABCD* rule for melanoma detection [31]. These features are certainly effective for the classification purpose, as seen from the performance of some classification-based systems in this domain, claiming a correct classification up to 100% [27] or specificity/sensitivity of 92.34%/93.33% [20]. However, features good for classification or distinguishing one disease from another may not be suitable for retrieval and display of similar appearing lesions. In this retrieval system, we are looking for similar images in term of colour, texture, shape, etc. By selecting and extracting good representative features, we may be able to identify images similar to an unknown query image, whether it belongs to the same disease group or not. Similar images belonging to different classes may give an idea about the certainty of classification.

Skin lesions appear mainly characterised by their colour and texture. In this section we will describe simple features that can capture such properties.

Colour Colour features are represented by the mean colour $\mu = (\mu_R, \mu_G, \mu_B)$ of the lesion and their covariance matrix Σ. Let

$$\mu_X = \frac{1}{N}\sum_{i=1}^{N} X_i \quad \text{and} \quad C_{XY} = \frac{1}{N}\left[\sum_{i=1}^{N} X_i Y_i\right] - \mu_X \mu_Y \tag{1}$$

where: N is the number of pixels in the lesion, X_i the colour component of channel X ($X, Y \in \{R, G, B\}$) of pixel i. Assuming to use the original RGB (Red, Green, Blue) colour space, the covariance matrix is:

$$\Sigma = \begin{bmatrix} C_{RR}\, C_{RG}\, C_{RB} \\ C_{GR}\, C_{GG}\, C_{GB} \\ C_{BR}\, C_{BG}\, C_{BB} \end{bmatrix} \tag{2}$$

In this work, RGB, HSV (Hue, Saturation, Value) and CIE_Lab, CIE_Lch (Munsell colour coordinate system [17]) and Otha [32] colour spaces are used.

A number of normalisation techniques have been applied before extracting colour features. We normalised each colour component by the average of the same component of the healthy skin of the same patient, because it had best performance.

After experimenting with the 5 different colour spaces, we choose the normalised RGB, because it gave slightly better results than the other colour spaces.

Texture Texture features are extracted from generalised co-occurrence matrices (GCM). Assume an image I having N_x columns, N_y rows and N_g grey levels. Let $L_x = \{1, 2, \cdots, N_x\}$ be the columns, $L_y = \{1, 2, \cdots, N_y\}$ be the rows, and $G_x = \{0, 1, \cdots, N_g - 1\}$ be the set of quantised grey levels. The co-occurrence matrix P_δ is a matrix of dimension $N_g \times N_g$, where [33]:

$$P_\delta(i, j) = \#\{((k, l), (m, n)) \in (L_y \times L_x) \times (L_y \times L_x) | I(k, l) = i, I(m, n) = j\} \tag{3}$$

i.e. the number of co-occurrences of the pair of grey level i and j which are a distance $\delta = (d, \theta)$ apart. In our work, the pixel pairs (k, l) and (m, n) have distance $d = 5, 10, 15, 20, 25, 30$ and orientation $\theta = 0°, 45°, 90°, 135°$, i.e. $(m = k + d, n = l), (m = k + d, n = l + d), (m = k, n = l + d), (m = k - d, n = l + d)$.

Generalised co-occurrence matrices are the extension of the co-occurrence matrix to multispectral images, i.e. images coded on n colour channels. Let u and v be two colour channels. The generalised co-occurrence matrices are:

$$P_\delta^{(u,v)}(i, j) = \#\{((k, l), (m, n)) \in (L_y \times L_x) \times (L_y \times L_x) | I_u(k, l) = i, I_v(m, n) = j\} \tag{4}$$

For example, in case of colour images, coded on three channels (RGB), we have six cooccurrence matrices: (RR),(GG),(BB) that are the same as grey level co-occurrence matrices computed on one channel and (RG), (RB), (GB) that take into account the correlations between the channels.

In order to have orientation invariance for our set of GCMs, we averaged the matrices with respect to θ. Quantisation levels $N_G = 64, 128, 256$ are used for the three colour spaces: RGB, HSV and CIE_Lab.

From each GCM we extracted 12 texture features: energy, contrast, correlation, entropy, homogeneity, inverse difference moment, cluster shade, cluster prominence, max probability, autocorrelation, dissimilarity and variance as defined in [33], for a total of 3888 texture features (12 features $\times$ 6 inter-pixel distances $\times$ 6 colour pairs $\times$ 3 colour spaces $\times$ 3 grey level quantisations).

These texture features are extracted from GCMs calculated over the lesion area of the image, as well as over a patch of healthy skin of the same image. Differences and ratios of each of these values are calculated:

$$feature_{l-s} = feature_{lesion} - feature_{healthy\_skin} \tag{5}$$

$$feature_{l/s} = feature_{lesion} / feature_{healthy\_skin} \tag{6}$$

3 Similarity matching

The retrieval system is based on a similarity measure defined between the query image Q and a database image I.

For colour covariance-based features, the Bhattacharyya distance metric is used as follow:

$$D_C(Q,I) = \frac{1}{8}(\mu_Q - \mu_I)^T \left[\frac{(\Sigma_Q + \Sigma_I)}{2}\right]^{-1} (\mu_Q - \mu_I) + \frac{1}{2}\ln \frac{\left|\frac{(\Sigma_Q+\Sigma_I)}{2}\right|}{\sqrt{|\Sigma_Q||\Sigma_I|}} \tag{7}$$

where μ_Q and μ_I are the average (over all pixels in the lesion) colour feature vectors, Σ_Q and Σ_I are the covariance matrices of the lesion of Q and I respectively, and $|\cdot|$ denotes the matrix determinant.

The Euclidean distance $D_T(Q,I)$ is used for distances between a subset of texture features f_{subset}, selected as described later.

$$D_T(Q,I) = \|f^Q_{subset} - f^I_{subset}\| = \sqrt{\sum_{i=1}^{m}(f^Q_i - f^I_i)^2} \tag{8}$$

Other metric distances have been considered, but gave worse results.

We aggregated the two distances into a similarity matching function as:

$$S(Q,I) = w_C \cdot D_C(Q,I) + (1 - w_C) \cdot D_T(Q,I) \tag{9}$$

where w_C is a weighting factor that has been selected experimentally, after trying all the values: $\{0.1, 0.2, \cdots, 0.9\}$. In our case, $w_C = 0.7$ gave the best results. A low value of S indicates a high similarity.

4 Hierarchical retrieval system

By observing the scatter plot of some of the best features selected during previous experiments [29], we noted that our image classes can be grouped in two

main groups. The first group, hence called *Group1*, contains lesion of classes: Actinic Keratosis (AK), Basal Cell Carcinoma (BCC) and Squamous Cell Carcinoma (SCC). The second group, hence called *Group2*, contains lesions of classes: Melanocytic Nevus / Mole (ML) and Seborrhoeic Keratosis (SK). It is worth to note that AK, BCC, SCC, ML and SK are diagnostic classes defined by dermatologist, while the two groups are constructed by clustering classes containing images which present visual similar characteristics.

However we can give some meaning to two groups observing that the fist group comprises BCC and SCC that are two common types of skin cancer and AK that is not cancer, but can be considered as a pre-malignant condition and it is visual similar to some of the BCC. In the second group ML and SK are both benign form of skin lesions.

The class grouping leads to the hierarchical structure shown in Fig. 1.

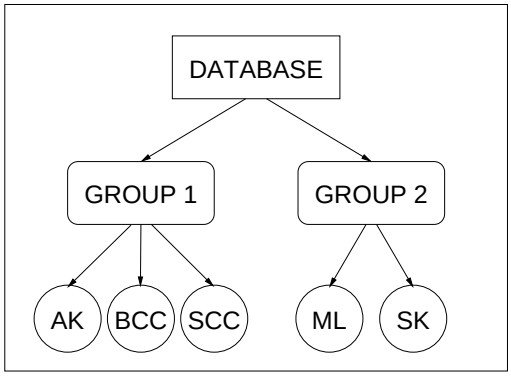

Fig. 1. Block diagram of the hierarchical organisation of skin lesion classes

This structure makes a coarse separation between classes at upper level while finer decisions are made at lower level. As a results, this scheme decomposes the original problem into 3 sub-problems.

Our retrieval system consists of two steps, each of them use a different set of features selected as described later.

We should also distinguish the learning phase, during which we select the best sets of features from the retrieval operative phase, when these sets of features are combined and used to retrieve the most similar images to the query image.

4.1 Learning phase

In the learning stage we perform the feature selection using three distinct systems. The first system uses as a database all the images and discriminates between the two groups, i.e. retrieves images that belong to the same group of the query image, irrespective of the class. For example, if the query image is a BCC, all the images in the diagnostic classes BCC, AK and SCC are considered

corrected retrieval answers. The other two systems use as database the images of one group each and discriminate between the classes. The first of these two systems use as database only the images of the first group (AK, BCC, SCC), and its goal is to retrieve image of the same class of the query image. Similarly the other system works on the image of the second group (ML and SK).

Feature selection is performed using a sequential forward selection (SFS) on the texture features described in Sec. 2. One could choose other search methods in the wrapper framework such as sequential backward elimination, forward-backward, sequential floating searches, or genetic algorithms [34]. SFS is a greedy search algorithm that adds one feature at a time. This method adds the feature that, when combined with the current chosen set, yields the largest improvement to our feature selection criterion. SFS does not guarantee an optimal solution, but it is simple and is sufficient for our purpose.

The wrapper framework is the retrieval system and its performance is used to guide the feature selection In all the three systems the objective is the maximisation of the number of correctly retrieved images, i.e. images belonging to the same group of the query image in the first system, and images belonging to the same class in the latter two systems. This measure is closely related to *precision*, that is the ratio of the number of relevant images returned to the total number of images returned. We averaged it using each image in the database (or in the sub-databases for the latter two systems) as query image, and asking the system to retrieve 10 similar images for each presented image (not retrieving itself). Often, in the information retrieval context, the F-measure, that is a weighted harmonic mean of precision and recall is often used. In our case, due to the uneven distribution of lesions into classes, it seemed more appropriate to maximise the precision.

At the end of the learning phase, three sets of features are selected. Let us call set_A the feature set of the first system, and set_{B1} and set_{B2} the feature sets of the latter two systems.

It can be argued that three systems need to be trained instead of one. It is worth to note that the training time needed for the feature selection of the latter two systems is reduced. Indeed these two systems are trained using only a subset of the images and therefore only the distances between the query image and the images in the subset need to be calculated. Moreover, since this is done in a off-line manner, the training time is not as critical as the optimality of the feature subset it generates.

As pointed out in Sec. 1, the use of a very large number of features not only increases the computational and memory requirements, preventing a real-time response during the operative phase of the system, but can also have negative effects on the convergence of the algorithm (due to the curse of dimensionality) [35]. The proposed methodology aims at solving this problem by selecting for each retrieval sub-system only the features that are relevant to that system and using that set of features in the corresponding step of the hierarchical system during the operative phase described in the following section.

4.2 Operative phase

During the operative phase we use the similarity matching function, defined in (9). We perform the retrieval in two steps and combine the three sets of features as follow.

Given a query image Q, first we determine $S_A(Q, I)$ between Q and each of the N images I in the database using the feature set_A to calculate the term D_T of eq. (9). Then we sort the images according to their similitude to the query image:

$$I_S = sorting(S_A(Q, I)) \tag{10}$$

At this point we need to decide which second set of features to use in the second step.

By observing to which group belong the majority of the first K sorted images we calculate a second similitude between the query image Q and each of the K images I' as follow:

$$S_B(Q, I') = \begin{cases} S_{B1}(Q, I')) \; if \; \#\{I' \in Group1\} > \#\{I' \in Group2\} \\ S_{B2}(Q, I')) \; otherwise \end{cases} \tag{11}$$

where $S_{B1}(Q, I')$ and $S_{B2}(Q, I')$ are calculated using set_{B1} and set_{B2}, respectively.

This decision is based on the assumption that if the majority of the first K images retrieved by the first step belong to a certain group, the query image (which group is unknown) should belong to that group. Therefore we need to search for similar images within that group, using the set of features that characterises that group and better discriminates the diagnostic classes within that group.

Then we resort the images I' according to their recalculated similitude to the query image:

$$I'_S = sorting(S_B(Q, I')) \tag{12}$$

Finally the images having the smallest S to the query image are presented as retrieval results.

A draw-back of the proposed method is that errors on the first step can not be corrected in the second step. Only the first K images retrieved by the first step will constitute the image database searched in the second step.

Moreover the choice of K is crucial: using a too large value of K we void the first step. On the other hand the use of a too small value of K risks to loose a high number of potential good retrieval results. In our experiments $K = N/10$ gave the best results.

5 Results and evaluation

Our image database comprises 533 lesions, belonging to 5 classes (20 AK, 116 BCC, 224 ML, 20 SCC, 153 SK). Images are acquired using a Canon EOS 350D SRL camera, having a resolution of about 0.03 mm. Lesions are segmented using the method described in [30]. The ground truth used for the experiments is based on agreed classifications by 2 dermatologists.

We compare our results with the results obtained using a non hierarchical approach, i.e. a flat retrieval system that uses a single set of features for all the 5 classes.

The standard parameters are used, namely the *precision* (number of correct retrieved images divided by total number of retrieved images) and the *recall* (number of correct retrieved images divided by total number of images of the same in the database). The effectiveness of the retrieval system is measured with the precision/recall curves that are commonly used in the information retrieval domain. A retrieved image is considered to be a correct match if it belongs to the same diagnostic class to which the query image belongs.

Fig. 2 shows two typical examples of image retrieval. On the left there are typical screenshots of our system. The improvement in retrieval results can be easily visualized by comparing the precision/scope plots on the right.

In Fig. 3, we show average precision/recall and precision/scope curves obtained by evaluating top N retrieved results (*scope*). Each image from the database is served as a query image. Results have been provided as averages obtained across all the queries.

It is clear that the proposed hierarchical retrieval system outperforms the non hierarchical system, with an improvement of the precision values between 4% and 5% for the first 10 retrieved images.

Moreover, as shown in Fig. 4, the feature set that best discriminates between the groups at the first level of the hierarchy is different from the feature sets that best discriminate at the second level, and these two sets also differ between each other. Hence, we prove that there is a need for a hierarchical scheme.

5.1 Influence of the K parameter

Fig. 5.1 shows the effect on retrieval performance at varying the number K of the first retrieved images that undergo through the second step of the system.

The optimal number seems to be at $K = 50$ that corresponds to about 10% of our database size. On the extreme, $K = N$ is the worst scenario, where the hierarchy has no effect, while for low values of K the precision increases at low values of scope, but its trend decreases pretty soon, meaning that only few good images are retrieved. Therefore we could advise to use low values of K if we are only interested in the very few top retrieved images.

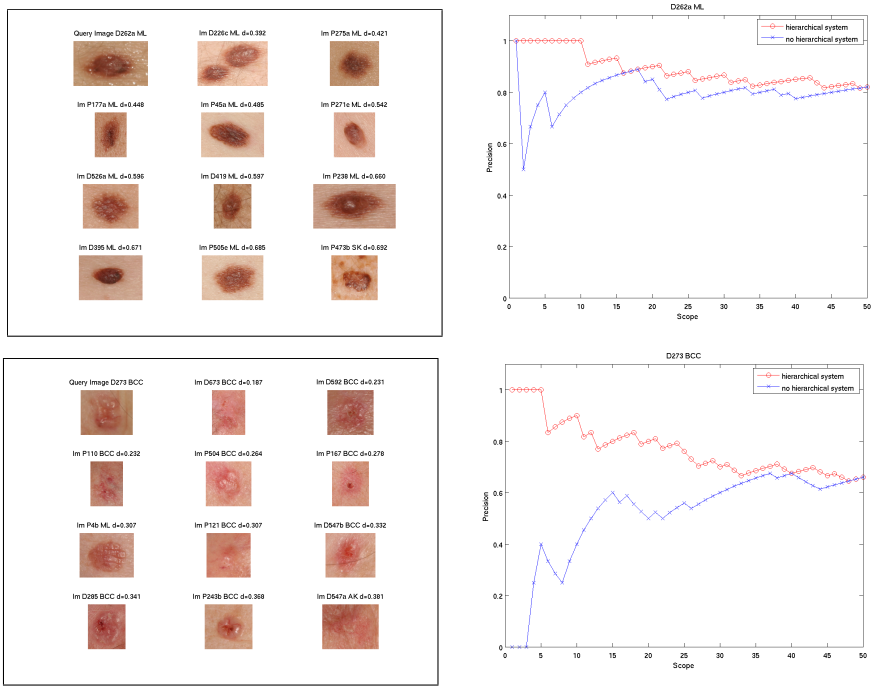

Fig. 2. Two typical screenshots showing retrieved images similar to the query image (top left image). Comparative precision/scope plots between the hierarchical (red) and the flat system (blue) relative to the query on the left

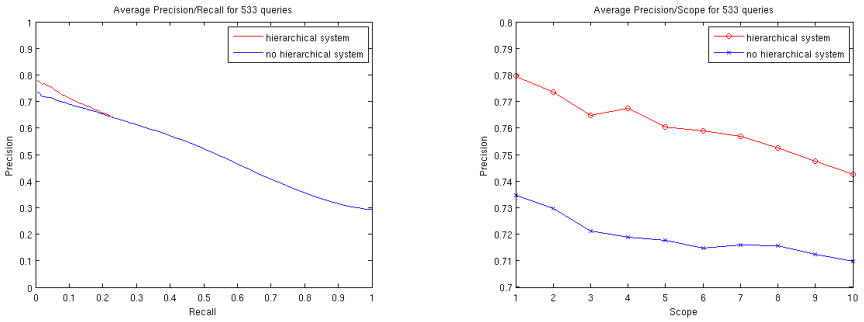

Fig. 3. Precision/Recall curves (left) and Precision/Scope curve (right) using a hierarchical (red curve) and a non hierarchical (blue curve) system.

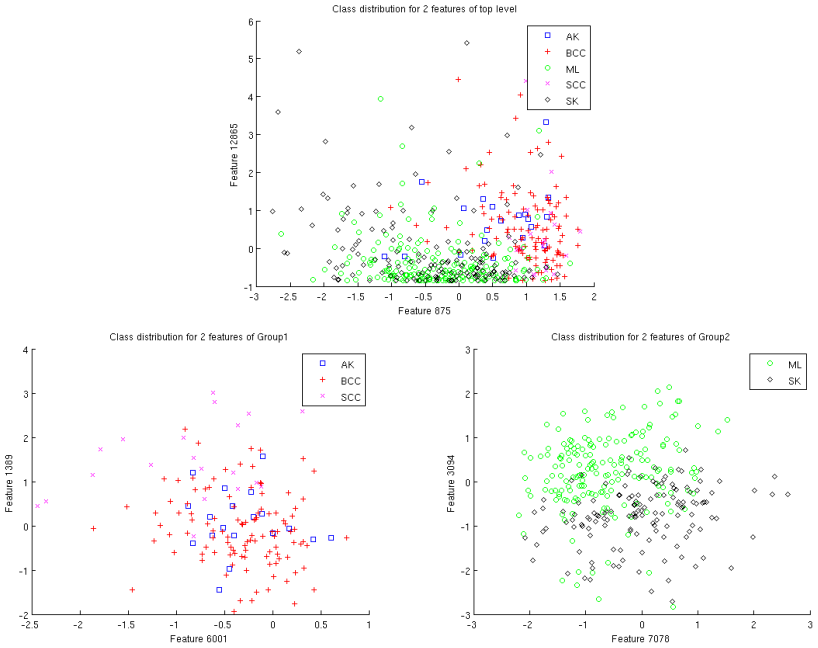

Fig. 4. Scatter plots of the some features of the three sets

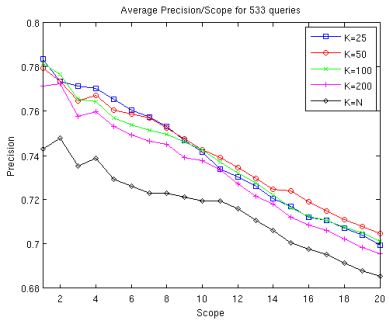

Fig. 5. Precision/Scope curves varying the value of K

6 Conclusions

We have presented a hierarchical CBIR system as a diagnostic aid for skin lesion images. We believe that retrieving and displaying images with known pathology that are visually similar to an image being evaluated may provide intuitive clinical decision support to dermatologists. The hierarchical structure of our system has proved to improves the performance of the system compared to the flat one.

Further studies will include the extraction of other texture-related features (i.e. fractal dimension, Gabor- and Tamura-based) as well as shape and boundary features. A larger database is being reached. We plan also to include relevance feedback, which is commonly used in image retrieval, but has not yet been used for medical images.

In the future, it would be interesting to apply our hierarchical approach on other domains and extending it to more than two hierarchical levels.

Acknowledgements: We thank the Wellcome Trust for funding this project.

References

1. Rui, Y., Huang, T.S., Chang, S.F.: Image retrieval: Current techniques, promin-sign directions, and open issues. Journal of Visual Communication and Image Representation **10** (1999) 39–62
2. Smeulders, A.W.M., Worring, M., Santini, S., Gupta, A., Jain, R.: Content-based image retrieval at the end of the early years. IEEE Transactions on Pattern Analysis and Machine Intelligence **22**(12) (2000) 1349–1380
3. Datta, R., Joshi, D., Li, J., Wang, J.Z.: Image retrieval: Ideas, influences, and trends of the new age. ACM Computing Surveys **40**(2) (April 2008) 5:1–5:60
4. Müller, H., Michoux, N., Bandon, D., Geissbuhler, A.: A review of content-based image retrieval systems in medical applications - clinical benefits and future directions. International Journal of Medical Informatics **73** (2004) 1–23
5. Guldogan, E., Gabbouj, M.: Feature selection for content-based image retrieval. Signal Image and Video Processing **2**(3) (September 2008) 241–250
6. Jain, A.K., Duin, R.P.W., Mao, J.: Statistical pattern recognition: A review. IEEE Transactions on Pattern Analysis and Machine Intelligence **22**(1) (2000) 4–37
7. Ribeiro, M.X., Bugatti, P.H., Traina, C., Marques, P.M.A., Rosa, N.A., Traina, A.J.M.: Supporting content-based image retrieval and computer-aided diagnosis systems with association rule-based techniques. Data & Knowledge Engineering **68**(12) (2009) 1370–1382
8. Bower, G.H., Clark, M.C., Lesgold, A.M., Winzen, D.: Hierarchical retrieval schemes in recall of categorized word lists. Journal fo Verbal Learning and Verbal Behavior **8** (1969) 323–343
9. Levinson, R., Ellis, G.: Multi-level hierarchical retrieval. In: In 6th Annual Conceptual Graphs Workshop. (1996) 285–310
10. Chow, S., Rahman, M., Wu, S.: Content-based image retrieval by using tree-structured features and multi-layer self-organizing map. Pattern Anal. Appl. **9**(1) (2006) 1–20
11. Lehmann, T.M., Güld, M.O., Thies, C., Fischer, B., Spitzer, K., Keysers, D., Ney, H., Kohnen, M., Schubert, H., Wein, B.: Content-based image retrieval in medical applications. Methods of Information in Medicine **43**(4) (2004) 354–361

12. Wichert, A.: Content-based image retrieval by hierarchical linear subspace method. Journal of Intelligent Information Systems **31**(1) (2008) 85–107
13. Dy, J., Brodley, C., Kak, A., Broderick, L., Aisen, A.: Unsupervised feature selection applied to content-based retrieval of lung images. Pattern Analysis and Machine Intelligence, IEEE Transactions on **25**(3) (mar. 2003) 373 – 378
14. Swets, D., Weng, J.: Hierarchical discriminant analysis for image retrieval. IEEE Transactions on Pattern Analysis and Machine Intelligence **21**(5) (may. 1999) 386 –401
15. Dermnet: the dermatologist's image resource (2007) Dermatology Image Altas, available at: http://www.dermnet.com/.
16. Cohen, B.A., Lehmann, C.U.: Dermatlas (2000-2009) Dermatology Image Altas, available at: http://dermatlas.med.jhmi.edu/derm/.
17. Rahman, M.M., Desai, B.C., Bhattacharya, P.: Image retrieval-based decision support system for dermatoscopic images. In: IEEE Symposium on Computer-Based Medical Systems, Los Alamitos, CA, USA, IEEE Computer Society (2006) 285–290
18. Chung, S.M., Wang, Q.: Content-based retrieval and data mining of a skin cancer image database. In: International Conference on Information Technology: Coding and Computing (ITCC 2001), Los Alamitos, CA, USA, IEEE Computer Society (2001) 611–615
19. Celebi, M.E., Aslandogan, Y.A.: Content-based image retrieval incorporating models of human perception. In: International Conference on Information Technology: Coding and Computing (ITCC 2004). Volume 2., Los Alamitos, CA, USA, IEEE Computer Society (2004) 241–245
20. Celebi, M.E., Kingravi, H.A., Uddin, B., Iyatomi, H., Aslandogan, Y.A., Stoecker, W.V., Moss, R.H.: A methodological approach to the classification of dermoscopy images. Computerized Medical Imaging and Graphics **31**(6) (2007) 362 – 373
21. Dorileo, E.A.G., Frade, M.A.C., Roselino, A.M.F., Rangayyan, R.M., Azevedo-Marques, P.M.: Color image processing and content-based image retrieval techniques for the analysis of dermatological lesions. 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS 2008) (August 2008) 1230–1233
22. Celebi, M.E., Iyatomi, H., Schaefer, G., Stoecker, W.V.: Lesion border detection in dermoscopy images. Computerized Medical Imaging and Graphics **33**(2) (2009) 148–153
23. Wollina, U., Burroni, M., Torricelli, R., Gilardi, S., Dell'Eva, G., Helm, C., Bardey, W.: Digital dermoscopy in clinical practise: a three-centre analysis. Skin Research and Technology **13** (May 2007) 133–142(10)
24. Seidenari, S., Pellacani, G., Pepe, P.: Digital videomicroscopy improves diagnostic accuracy for melanoma. Journal of the American Academy of Dermatology **39**(2) (1998) 175–181
25. Lee, T.K., Claridge, E.: Predictive power of irregular border shapes for malignant melanomas. Skin Research and Technology **11**(1) (2005) 1–8
26. Schmid-Saugeons, P., Guillod, J., Thiran, J.P.: Towards a computer-aided diagnosis system for pigmented skin lesions. Computerized Medical Imaging and Graphics **27** (2003) 65–78
27. Maglogiannis, I., Pavlopoulos, S., Koutsouris, D.: An integrated computer supported acquisition, handling, and characterization system for pigmented skin lesions in dermatological images. IEEE Transactions on Information Technology in Biomedicine **9**(1) (2005) 86–98

28. Maglogiannis, I., Doukas, C.N.: Overview of advanced computer vision systems for skin lesions characterization. IEEE Transaction on Information Technology in Biomedicine **13**(5) (2009) 721–733
29. Ballerini, L., Li, X., Fisher, R.B., Rees, J.: A query-by-example content-based image retrieval system of non-melanoma skin lesions. In Caputo, B., ed.: Proc. MICCAI-09 Workshop MCBR-CDS 2009: Medical Content-based Retrieval for Clinical Decision Support. Number 5853 in LNCS, Springer-Verlag Berlin Heidelberg (2010) 31–38
30. Xiang, L., Aldridge, B., Ballerini, L., Fisher, R., Rees, J.: Depth data improves skin lesion segmentation. In et al., G.Z.Y., ed.: Proc. 12th Int. Conf. on Medical Image Computing and Computer Assisted Intervention (MICCAI), London. (2009) 1100–1107
31. Johr, R.H.: Dermoscopy: alternative melanocytic algorithms–the ABCD rule of dermatoscopy, menzies scoring method, and 7-point checklist. Clinics in Dermatology **20**(3) (May-June 2002) 240–247
32. Ohta, Y.I., Kanade, T., Sakai, T.: Color information for region segmentation. Computer Graphics and Image Processing **13**(1) (July 1980) 222 – 241
33. Haralick, R.M., Shanmungam, K., Dinstein, I.: Textural features for image classification. IEEE Transactions on Systems, Man and Cybernetics **3**(6) (1973) 610–621
34. Goldberg, D.E.: Genetic Algorithms in Search, Optimization, and Machine Learning. Addison-Wesley, Reading, MA (1989)
35. Grigorova, A., De Natale, F., Dagli, C., Huang, T.: Content-based image retrieval by feature adaptation and relevance feedback. IEEE Transactions on Multimedia **9**(6) (2007) 1183–1192